PCT/GB05/05526 #2

Rec'd PCT/PTO 17 JUN 2005

The Patent Office

INVESTOR IN PEOPLE

The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ

PRIORITY DOCUMENT

SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

REC'D 01 MAR 2004
WIPO PCT

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

Signed [illegible]

Dated 13 February 2004

BEST AVAILABLE COPY

Patents Form 1/77

Patents Act 1977
(Rule 16)

THE PATENT OFFICE
A
18 DEC 2002
RECEIVED BY FAX

The Patent Office
PATENTS • DESIGNS • COPYRIGHT • TRADE MARKS

EC02 E771634-1 D02884
P01/7700 0.00-0229354.6

Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)

The Patent Office
Cardiff Road
Newport
South Wales
NP10 8QQ

1. Your reference — P33024-/TSI/ILI/KJO

2. Patent application number *(The Patent Office will fill in this part)* — 0229354.6 — 18 DEC 2002

3. Full name, address and postcode of the or of each applicant *(underline all surnames)*

   The Robert Gordon University
   Schoolhill
   Aberdeen
   AB10 1FR

   Patents ADP number *(if you know it)* — 6622997001

   If the applicant is a corporate body, give the country/state of its incorporation — United Kingdom

4. Title of the invention — "Video Encoding"

5. Name of your agent *(if you have one)* — Murgitroyd & Company

   "Address for service" in the United Kingdom to which all correspondence should be sent *(including the postcode)*

   Scotland House
   165-169 Scotland Street
   Glasgow
   G5 8PL

   Patents ADP number *(if you know it)* — 1198015

6. If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and *(if you know it)* the or each application number

Country	Priority application number *(if you know it)*	Date of filing *(day / month / year)*

7. If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application

Number of earlier application	Date of filing *(day / month / year)*

8. Is a statement of inventorship and of right to grant of a patent required in support of this request? *(Answer 'Yes' if:* — Yes

   a) any applicant named in part 3 is not an inventor, or
   b) there is an inventor who is not named as an applicant, or
   c) any named applicant is a corporate body.
   See note (d))

Patents Form 1/77

9. Enter the number of sheets for any of the following items you are filing with this form. Do not count copies of the same document

Continuation sheets of this form	-
Description	12
Claim(s)	-
Abstract	-
Drawing(s)	3

10. If you are also filing any of the following, state how many against each item.

Priority documents	-
Translations of priority documents	-
Statement of inventorship and right to grant of a patent *(Patents Form 7/77)*	-
Request for preliminary examination and search *(Patents Form 9/77)*	-
Request for substantive examination *(Patents Form 10/77)*	-
Any other documents *(please specify)*	-

11. I/We request the grant of a patent on the basis of this application.

Signature Murgitroyd & Company — Murgitroyd & Company

Date 18 December 2002

12. Name and daytime telephone number of person to contact in the United Kingdom — KEITH JONES — 0141 307 8400

Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

Notes

a) *If you need help to fill in this form or you have any questions, please contact the Patent Office on 08459 500505.*

b) *Write your answers in capital letters using black ink or you may type them.*

c) *If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.*

d) *If you have answered 'Yes' Patents Form 7/77 will need to be filed.*

e) *Once you have filled in the form you must remember to sign and date it.*

f) *For details of the fee and ways to pay please contact the Patent Office.*

Video Encoding

The invention relates to video encoders and in particular to reducing the computational complexity when encoding video.

Video encoders and decoders (CODECs) based on video encoding standards such as H263 and MPEG-4 are well known in the art of video compression.

The development of these standards has led to the ability to send video over much smaller bandwidths with only a minor reduction in quality. However, decoding and, more specifically, encoding, requires a significant amount of computational processing resources. For mobile devices, such as personal digital assistants (PDA's) or mobile telephones, power usage is closely related to processor utilisation and therefore relates to the life of the battery charge. It is obviously desirable to reduce the amount of processing in mobile devices to

the amount of processing in mobile devices to increase the operable time of the device for each battery charge. In general-purpose personal computers, CODECs must share processing resources with other applications. This has contributed to the drive to reduce processing utilisation, and therefore power drain, without compromising viewing quality.

In many video applications, such as tele-conferences, the majority of the area captured by the camera is static. In these cases, power resources or processor resources are being used unnecessarily to encode areas which have not changed significantly from a reference video frame.

The typical steps required to process the pictures in a video by an encoder such as one that is H263 or MPEG-4 Simple Profile compatible, are described as an example.

The first step requires that reference pictures be selected for the current picture. These reference pictures are divided into non-overlapping macroblocks. Each macroblock comprises four luminance blocks and two chrominance blocks, each block comprising 8 pixels by 8 pixels.

It is well known that the steps in the encoding process that typically require the greatest computational time are the motion estimation, the

forward discrete cosine transform (FDCT) and the inverse discrete cosine transform (IDCT).

The motion estimation step looks for similarities between the current picture and one or more reference pictures. For each macroblock in the current picture, a search is carried out to identify a prediction macroblock in the reference picture which best matches the current macroblock in the current picture. The prediction macroblock is identified by a motion vector (MV) which indicates a distance offset from the current macroblock. The prediction macroblock is then subtracted from the current macroblock to form a prediction error (PE) macroblock. This PE macroblock is then discrete cosine transformed, which transforms an image from the spatial domain to the frequency domain and outputs a matrix of coefficients relating to the spectral sub-bands. For most pictures much of the signal energy is at low frequencies, which is what the human eye is most sensitive to. The formed DCT matrix is then quantised which involves dividing the DCT coefficients by a quantizer value and then rounding to the nearest integer. This has the effect of reducing many of the higher frequency coefficients to zeros and is the step that will cause distortion to the image. Typically, the higher the quantizer step size, the poorer the quality of the image. The values from the matrix after the quantizer step are then re-ordered by "zigzag" scanning. This involves reading the values from the top left-hand corner of the matrix diagonally back

and forward down to the bottom right-hand corner of the matrix. This tends to group the zeros together which allows the stream to be efficiently run-level encoded (RLE) before eventually being converted into a bitstream by entropy encoding. Other "header" data is usually added at this point.

If the MV is equal to zero and the quantised DCT coefficients are all equal to zero then there is no need to include encoded data for the macroblock in the encoded bitstream. Instead, header information is included to indicate that the macroblock has been "skipped".

US 6,192,148 discloses a method for predicting whether a macroblock should be skipped prior to the DCT steps of the encoding process. This method decides whether to complete the steps after the motion estimation if the MV has been returned as zero, the mean absolute difference of the luminance values of the macroblock is less than a first threshold and the mean absolute difference of the chrominance values of the macroblock is less than a second threshold.

For the total encoding process the motion estimation and the FDCT and IDCT are typically the most processor intensive. The prior art only predicts skipped blocks after the step of motion estimation and therefore still contains a step in the process that can be considered processor intensive.

The present invention discloses a method to predict skipped macroblocks that requires no motion estimation or DCT steps.

According to one aspect, the invention provides a method of encoding video pictures comprising the steps of:

dividing the picture into regions;

predicting whether each region requires processing through further steps by comparing each region with a reference region. Hence, the invention avoids unnecessary use of resources by avoiding processor intensive operations where possible.

The further steps preferably include motion estimation and/or discrete cosine transform steps.

A region is preferably a non-overlapping macroblock.

A macroblock is preferably a sixteen by sixteen matrix of pixels.

Further preferably, a reference region is one or more macroblocks in the same position in the video picture but from one or more different reference time frames as selected by other encoding steps.

Preferably, the step of predicting includes two or more sub-steps.

Preferably, the sub-steps of the predicting step are calculations.

Preferably, one of the calculations is whether an estimate of the energy of some or all pixel values of the macroblock, optionally divided by the quantizer step size, is less than a predetermined threshold value.

Alternatively or further preferably, one of the calculations is whether an estimate of the values of certain discrete cosine transform coefficients for one or more sub-blocks of the macroblock, is less than a second threshold value.

Further preferably, the method of encoding pictures may be performed by a computer program embodied on a computer usable medium.

Further preferably, the method of encoding pictures may be performed by electronic circuitry.

The estimate of the values of certain discrete cosine transform coefficients may involve:
dividing the sub-blocks into four equal regions;
calculating the sum of absolute differences of the residual pixel values for each region of the sub-block, where the residual pixel value is the corresponding reference pixel luminance value subtracted from the current pixel luminance value;
estimating the low frequency discrete cosine transform coefficients for each region of the sub-blocks, such that:

$$Y_{01} = abs(A+C-B-D)$$
$$Y_{10} = abs(A+B-C-D)$$
$$Y_{11} = abs(A+D-B-C)$$

where Y_{01}, Y_{10} and Y_{11} represent the estimations of three low frequency discrete cosine transform coefficients and A, B, C and D represent the sum of absolute differences of each of the regions of the sub-block where A is the top left hand corner, B is the top right hand corner, C is the bottom left hand corner and D is the bottom right hand corner; and

selecting the maximum value of the estimate of the discrete cosine transform coefficients from all the estimates calculated.

The invention will now be described, by way of example, with reference to the figures of the drawings in which:

Figure 1 shows a flow diagram of a video picture encoding process.

Figure 2 shows a flow diagram of a macroblock encoding process

Figure 3 shows a flow diagram of a prediction decision process

With reference to Figure.1, a first step 102 reads a picture frame in a video sequence and divides it into non-overlapping macroblocks (MBs). Each MB comprises four luminance blocks and two chrominance

blocks, each block comprising 8 pixels by 8 pixels. Step 104 encodes the MB as shown in Figure 2.

With reference to Figure 2, a MB encoding process is shown 104, where a decision step 202 is performed before any other step.

The current H263 encoding process currently teaches that each MB in the video encoding process typically goes through the steps 204 to 226 or equivalent processes, in the order shown in Figure 2 or in a different order. Motion estimation step 204 identifies one or more prediction MB(s) each of which is defined by a MV indicating a distance offset from the current MB and a selection of a reference picture. Motion compensation step 206 subtracts the prediction MB from the current MB to form a Prediction Error (PE) MB. If the value of MV requires to be encoded (step 208), then MV is entropy encoded (step 210) optionally with reference to a predicted MV.

Each block of the PE MB is then forward discrete cosine transformed (FDCT) 212 which outputs a block of coefficients representing the spectral sub-bands of each of the PE blocks. The coefficients of the FDCT block are then quantized (for example through division by a quantizer step size) 214 and then rounded to the nearest integer. This has the effect of reducing many of the coefficients to zero. If there are any non-zero quantized coefficients

(Qcoeff) 216 then the resulting block is entropy encoded by steps 218 to 222.
In order to form a reconstructed picture for further predictions, the quantized coefficients (QCoeff) are re-scaled (for example by multiplication by a quantizer step size) 224 and transformed with an inverse discrete cosine transform (IDCT) 226. After the IDCT the reconstructed PE MB is added to the reference MB and stored for further prediction.

The decision step 228 looks at the output of the prior processes and if the MV is equal to zero and all the Qcoeffs are zero then the encoded information is not written to the bitstream but a skip MB indication is written instead. This means that all the processing time that has been used to encode the MB has not been necessary because the MB is regarded as similar to or the same as the previous MB.

Decision step 202 predicts whether the current MB is likely to be skipped, that is that after the process steps 202 - 226, the MB is not coded but a skip indication is written instead. If the Decision step 202 does predict that the MB would be skipped the MB is not passed on to the step 204 and the following process steps but skip information is passed directly to step 232.

With reference to Figure 3, a flow diagram is shown of the decision to skip the MB 202.

MBs that are skipped have zero MV and QCoeff. Both of these conditions are likely to be met if there is a strong similarity between the current MB and the same MB position in the reference frame. The energy of a residual MB formed by subtracting the reference MB, without motion compensation, from the current MB is approximated by the sum of absolute differences for the luminance part of the MB with zero displacement ($SAD0_{MB}$) given by:

$$SAD0_{MB} = \sum_{i=0}^{15} \sum_{j=0}^{15} \left| C_C(i,j) - C_P(i,j) \right| \qquad \textbf{Equation 1}$$

$C_C(i,j)$ and $C_p(i,j)$ are luminance samples from an MB in the current frame and in the same position in the reference frame, respectively.

The relationship between $SAD0_{MB}$ and the probability that the MB will be skipped also depends on the quantizer step size since a higher step size typically results in an increased proportion of skipped MBs.

A comparison of the calculation $SAD0_{MB}$ (optionally divided by the quantizer step size (Q)) 302 to a first threshold value gives a first comparison step 304. If the calculated value is greater than a first threshold value then the MB is passed to step 204 and enters a normal encoding process. If the calculated value is less than a first threshold value then a second calculation is performed 306.

Step 306 performs additional calculations on the residual MB. Each 8x8 luminance block is divided into four 4x4 blocks. A, B, C and D (Equation 2) are

the SAD values of each 4x4 block and R(i, j) are the residual pixel values without motion compensation.

$$A = \sum_{i=0}^{3} \sum_{j=0}^{3} |R(i,j)| \qquad B = \sum_{i=0}^{3} \sum_{j=3}^{7} |R(i,j)|$$

Equation 2

$$C = \sum_{i=4}^{7} \sum_{j=0}^{3} |R(i,j)| \qquad D = \sum_{i=4}^{7} \sum_{j=4}^{7} |R(i,j)|$$

Y_{01}, Y_{10} and Y_{11} (Equation 3) provide a low-complexity estimate of the magnitudes of the three low frequency DCT coefficients coeff(0,1), coeff(1,0) and coeff(1,1) respectively. If any of these coefficients is large then there is a high probability that the MB should not be skipped. $Y4\times4_{block}$ (Equation 4) is therefore used to predict whether each block may be skipped. The maximum for the luminance part of a macroblock is calculated using Equation 5.

$$Y_{01} = abs(A + C - B - D) \qquad Y_{10} = abs(A + B - C - D)$$

$$Y_{11} = abs(A + D - B - C)$$

Equation 3

$$Y4\times4_{Block} = MAX(Y_{01}, Y_{10}, Y_{11})$$

Equation 4

$$Y4\times4_{max} = MAX(Y4\times4_{block1}, Y4\times4_{block2}, Y4\times4_{block3}, Y4\times4_{block4})$$

Equation 5

The calculated value of $Y4{\times}4_{max}$ is compared with a second threshold 308. If the calculated value is less than a second threshold then the MB is skipped and the next step in the process is 232. If the calculated value is greater than a second threshold then the MB is passed to step 204 and the subsequent steps for encoding.

These steps typically have very little impact on computational complexity. $SAD0_{MB}$ is normally computed in the first step of any motion estimation algorithm and so there is no extra calculation required. Furthermore, the SAD values of each 4x4 block (A, B, C and D in Equation 2) may be calculated without penalty if $SAD0_{MB}$ is calculated by adding together the values of SAD for each 4x4-sample sub-block in the MB.

The additional computational requirements of the classification algorithm are the operations in Equations 3, 4 and 5 and these are typically not computationally intensive.

FIGURE 1

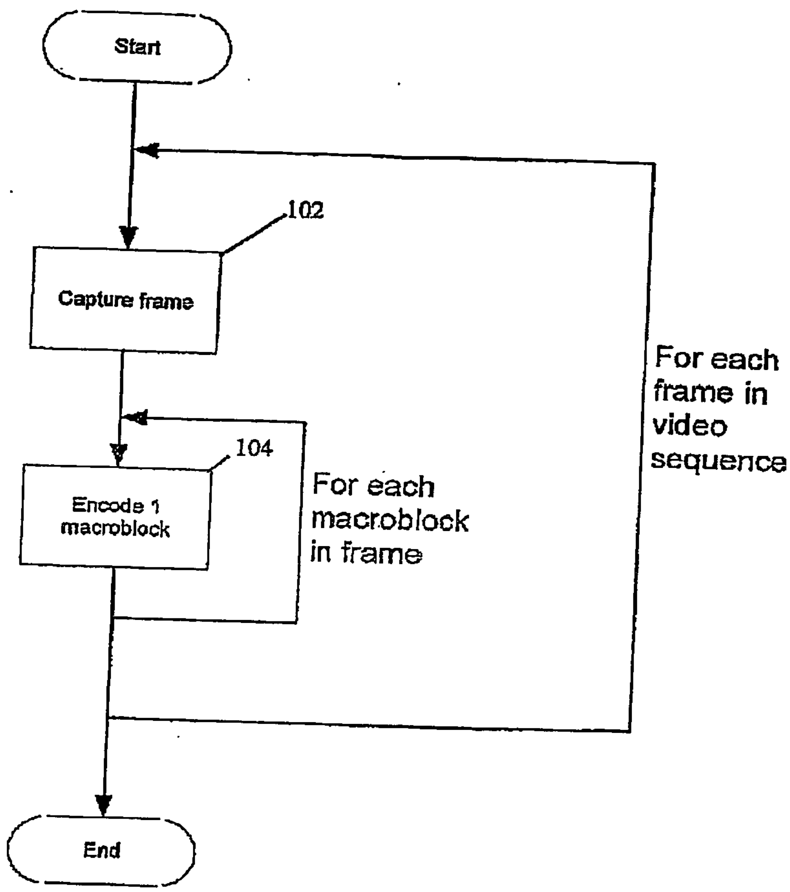

Start
102
Capture frame
104
Encode 1 macroblock
For each macroblock in frame
For each frame in video sequence
End

FIGURE 2

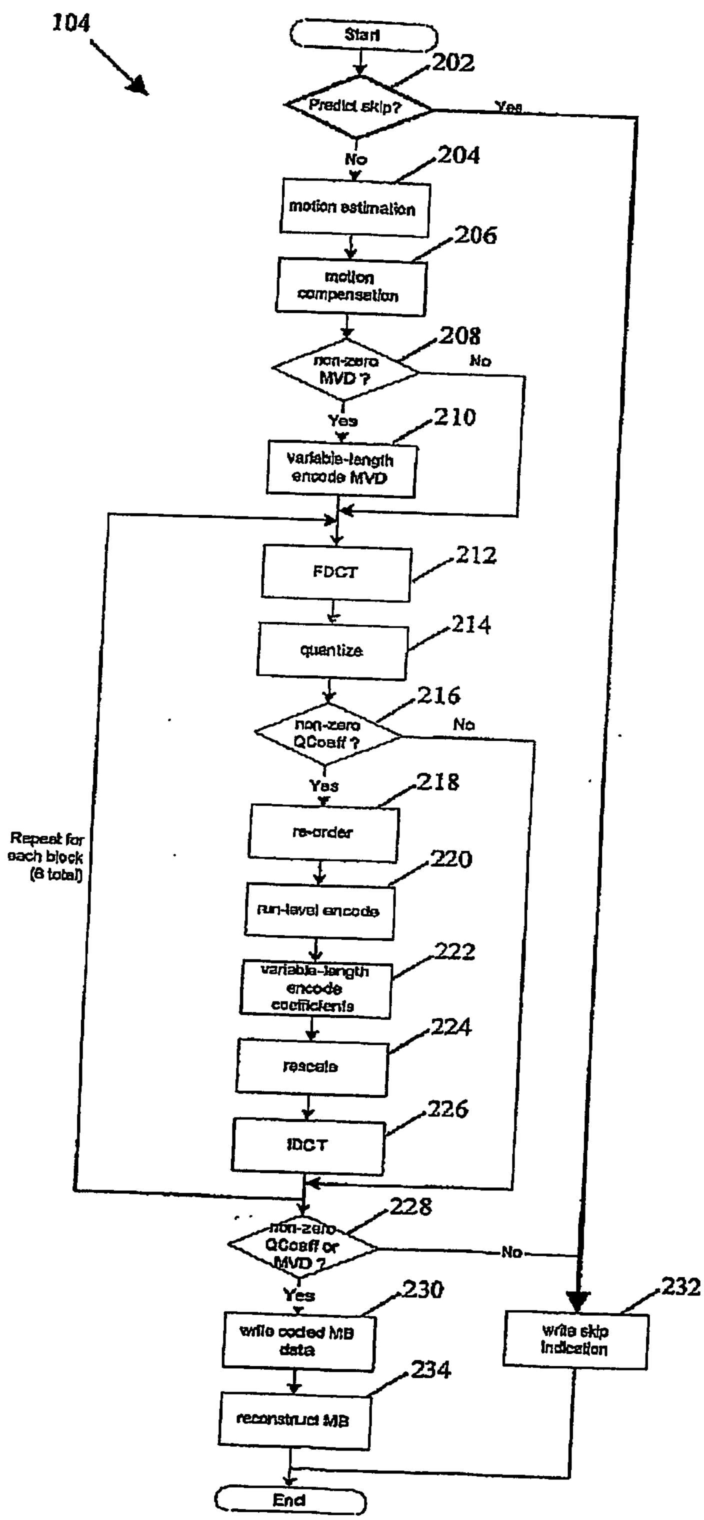
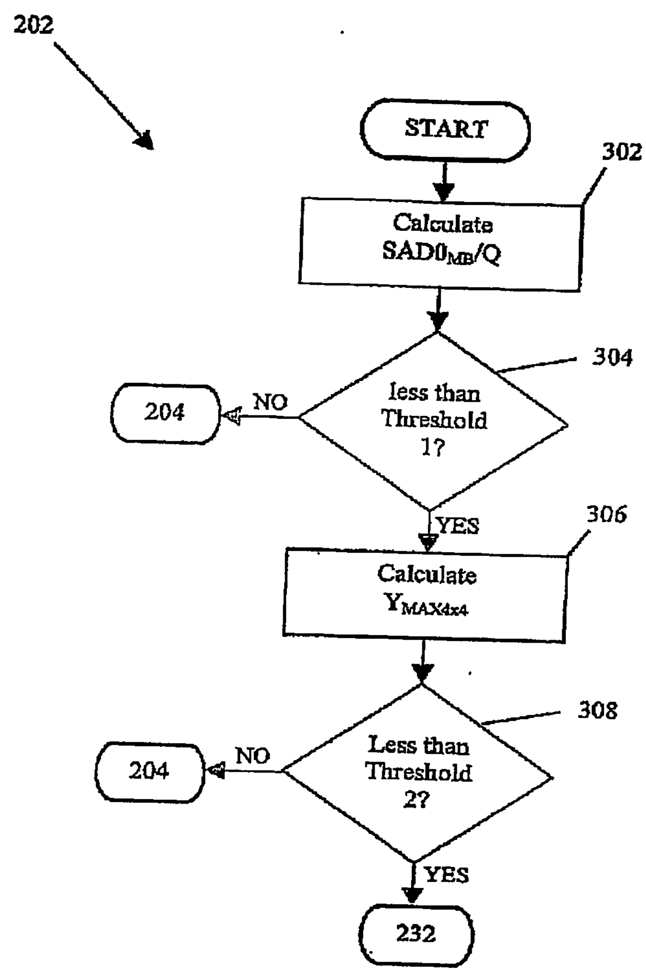

104
Start
202
Predict skip?
Yes
No
204
motion estimation
206
motion compensation
208
non-zero MVD ?
No
Yes
210
variable-length encode MVD
212
FDCT
214
quantize
216
non-zero QCoeff ?
No
Yes
218
re-order
220
run-level encode
222
variable-length encode coefficients
224
rescale
226
IDCT
Repeat for each block (6 total)
228
non-zero QCoeff or MVD ?
No
Yes
230
write coded MB data
232
write skip indication
234
reconstruct MB
End

FIGURE 3



202
START
302
Calculate
SAD0MB/Q
304
less than
Threshold
1?
NO
204
YES
306
Calculate
YMAX4x4
308
Less than
Threshold
2?
NO
204
YES
232